USING GRAMMAR IN ASSESSING ACADEMIC WRITING: A NARRATIVE STUDY OF EFL STUDENT EXPERIENCES

Sam Hermansyah[1], Muhammad Aksan[2], Nur Hikmah[3],Nurul Faradilah[4],Isumarni[5]

[12345]Universitas Muhammadiyah Sidenreng Rappang

Abstract:

Advances in technology have made it easy for all writers to check their academic writing using automated writing assessment programs. However, using this program can have both advantages and disadvantages. Therefore, its advantages and disadvantages should be considered. To address this need, this study aimed to identify the strengths and weaknesses of the Grammarly program as an automated writing assessment program for assessing academic writing. Using narrative research to explore the experiences of three Indonesian graduate students through interviews and documents, the program offers helpful color-coded feedback with explanations and examples, easy account access, rapid assessment speed, and It was indicated that they provide free services for academic evaluation. write. However, we have identified several concerns when using this program, including: Weaknesses of the program included some misleading feedback, weaknesses in English type recognition and reference lists, and lack of context and content assessment.

Further research is needed on the effectiveness of the feedback provided by Grammarly in improving the quality of students' writing.

Key Words: academic writing; automated writing evaluation; experience; Grammarly

INTRODUCTION

As most international journals and international conferences require the paper or article to be written in English, it is crucial for students, especially in English as a foreign language (henceforth, EFL) learning environment, to pay more attention to the mastery of English in their writing. Thus, to make the writing more clear and understandable, without creating an ambiguous meaning and misunderstanding information, a writing evaluation is required to be conducted before submitting a piece of writing to journal or conference.Since the technology is rapidly developing, many programs have been built to support the language learner in achieving better language proficiency. As an example, Automated Writing Evaluation (henceforth, AWE), as one of the products of technology, brings an innovation in evaluating writing. AWE is a computerized-program which can evaluate and rate writing automatically using online checking system. It uses an artificial intelligent developed by computational linguistics to rate and to score the writing submitted to the program (Ferster, et al., 2012; Wilson, 2016; Wilson, & Andrada, 2016) by analyzing the writing on lexical, syntactic, discourse, and grammar levels (Chen & Cheng, 2008; Chou, Moslehpour, & Yang, 2016) and provide diagnostic feedback and correction for the user (Chen & Cheng, 2008). Thus, the user can preview their evaluation result by looking at the feedback and correction given by the system and can start to revise the writing based on the evaluation given by themselves (Chen & Cheng, 2008; Ferster, et al., 2012) and save their time in checking and evaluating the writings (Chou, et al., 2016; Cotos, 2011; Roscoe, Wilson, Johnson, & Mayra, 2017). Equipped with the diagnostic feature, AWE program can be considered as an effective tool in evaluating writing.

However, not all AWE programs provide these features as benefits for the user and even, some AWE program leads their user to confusion and disappointment. Some scholars have been investigated some AWE program and revealed its strengths and weaknesses. Like *Criterion*, one model of successful AWE program development which brings satisfaction for their user on feedback clarity, but the scoring system can be deceived by omitting the error and writing a longer essay (Ebyary & Windeatt, 2010; Wang, 2013). *Intelligent Academic Discourse Evaluator* program can provide a clear feedback for the user and make them focus on revising their writing (Cotos, 2011), meanwhile, *ETIPS* only brings inaccurate score measurement in evaluating the writing which results in confusion (Scharber, Dexter, & Riedel, 2008). On the other hand, *Writing RoadmapTM 2.0* program can improve writing proficiency, but fails in providing comprehensive feedback and suggestion in revising the idea, content, and organization of the writing (Wang & Wang, 2012). Thus, the strengths and weaknesses

found on the program utilization still need to be counted as an essential consideration in deciding the implementation of AWE program to be used for evaluating writing.

On the other hand, *Grammarly*, as one of developed AWE program in 21st century, has been claimed as an easy tool that can help students and academies on their writing by checking the spelling, grammar, and punctuation errors and providing a comprehensive and useful feedback including correction and suggestion to make the writing more readable, clear, precise, effective, mistake-free, and impactful with a high rate of accuracy and evaluation speed (Grammarly, 2017). Nevertheless, there is a limited number of studies yet taking its claim into account and evaluating the process of evaluation given by *Grammarly* program. Thus, the user's view needs to be taken into account whether the utilization of this program brings the benefits as it claims or even leads the user to face any difficulties or confusion on the evaluation given.

Reflecting from this issue, I was interested in exploring the *Grammarly* user's experience on using this program. Taking three EFL postgraduate students' experience in evaluating their academic writing as *Grammarly* user, this study aimed to investigate the strengths and weaknesses of using *Grammarly* program in evaluating academic writing.

METHOD

To find the strengths and the weaknesses of *Grammarly* program, I have conducted a narrative inquiry research with a focus on exploring "a personal experience story" (Creswell, 2012) on utilizing AWE program in evaluating academic writing. This research design allowed me to explore an individual's personal experience deeply in a certain episode or situation (Denzin, 1989, as cited in Creswell, 2012); in this study, it explored the AWE program users' experiences on using this program in evaluating their academic writing. The experiences explored are related to any benefits and drawbacks they have faced along with this program utilization which can reveal the program's

strengths and weaknesses. To be more specific, the AWE program investigated in this study is *Grammarly* program. In utilizing their service, *Grammarly* offers two ways of checking process; free checking and premium checking. In Free-*Grammarly*, it corrects 150 types of errors, including critical grammar and spelling errors, and also checks any online writing, in email and social media. Meanwhile, in Premium-*Grammarly*, this program offers upgraded software to checks for over 400 checks and features, including vocabulary enhancement suggestions, plagiarism detection, and citation suggestions. Then, in this study, the free version *Grammarly* software is taken to be investigated since the users participated in this study only has the Free-*Grammarly* access.

The participants in this study are the *Grammarly* program users with different length of experience. These users were three Indonesian EFL postgraduate students; one of them is male and two others are females. Pseudonyms are given to these students, namely Mawar and Melati for the female students, and Bintang for the male one. They were purposefully selected since they have a long-term experiences on using this program in evaluating their academic writing, including assignments, conference papers, and articles for journal; Mawar has been using this automated writing evaluation program for seven months, Melati for one year and a half, and Bintang for two years and a half. In addition, these students also hold a Bachelor degree in English Education and now are taking Master degree in English Education. Thus, they have acquired some learning on English grammar and understand the basic of English grammar. Taking these students as the participants, they were expected to share their experiences on the *Grammarly* utilization in evaluating their academic writing.

Data and Source of Data

In conducting the study, I took qualitative data on the students' experiences. The qualitative data is in form of their answer on the interview conducted with these students. Then, a further justification of the data was also gained by taking the documentation of these students' *Grammarly* account.

Data Collecting Technique

As a narrative researcher, I learned about their experiences using a variety of narrative approaches, including document collection on their *Grammarly* account and interviews with them. They were firstly contacted to participate in the study and informed on their consent in participating in the study, including the interview and *Grammarly* account access. Then, a meeting for the interview was scheduled for each of them individually at different times. During the interview, a laptop was provided in

order to be able to access their account and also to help these students to recall their experience on using this program clearer. Their voices were recorded and transcribed, and the results of transcriptions were sent back to each of them, so they could confirm back of what they had said in the interview and could give further clarification from what they had been stated in the transcription. After being confirmed back, the final transcriptions were then analyzed.

Data Analysis Technique

After the transcriptions of the interview are being confirmed by the students, the final transcriptions were then analyzed. The students' answers were coded and arranged into certain them; covering the strengths and weaknesses of using this program in evaluating their academic writing. Then, a further interpretation was also created based on the themes found in their stories. The result of interpretation was then also sent back to these students to confirm the clarity of interpretation made supported with the documentation of their *Grammarly* account.

Result

From the result of interview conducted, these students have shared their experiences on utilizing *Grammarly* in evaluating their academic writing. From their stories, several strengths and weaknesses of this program were revealed.

1. Strengths

These students has experienced some benefits of utilizing this program which bcome the strengths of this program in evaluating academic writing, including useful feedback for learning gain, ease of access, high rate of evaluation speed, and free-service.

a. Useful Feedback for Learning Gain

The first strength revealed when these students talked about their stories on the utilization of *Grammarly* in evaluating our academic writing is the useful feedback given by this program which makes them learn better.

On Mawar experience, she admitted that she gained grammar learning by evaluating her papers using *Grammarly*. From "the clear feedback" given by this automated writing evaluation program, she could do self-revision and learn not to make the same errors for the next writing she wrote. She mentioned that this self-revision made her "realize on the writing errors made" and turned her "to be more careful in writing". She then also confirmed that the clear feedback given by *Grammarly* helped her in "learning from the mistake". She could grasp the explanation in the feedback easily since "it is completed with the example of the same error cases" which can help her to gain a better understanding of the writing context.

On the other hand, Melati also admitted her weaknesses in writing and told me how "*Grammarly* could help her improve her writing better by giving explanation and example". She pointed out that "the brief description given is easy to be caught". With a clear explanation and example given, "which is better than giving abstract explanation only", she seemed to be more aware of her writing errors and can learn how to write better.

Then, Bintang also admitted that using *Grammarly* in evaluating his academic writing gave him "a valuable learning experience", especially in enriching the English grammar. He also shared his experience on the feedback he received and he seemed to be satisfied with the feedback given. He saw it was effective in "giving him the explanation on his error and he could understand his error better". He also admitted that the evaluation helped him in determining the errors by "giving color-coded and cross on the errors", and this correction made him "aware of his errors in writing".

It can be seen that using *Grammarly* in evaluating academic writing seemed to bring a positive contribution for these students' writing improvement. As these students can learn from their mistakes, the use of this program could raise their awareness on their errors and give them a better understanding on revising their errors and preventing them from making similar errors.

b. Ease of Access in Downloading and Utilization

The ease of access becomes the second benefit felt in using *Grammarly*. The simple procedures given to create an account, to submit writing, and to evaluate writing were the major concern on its ease of access. In the interview, Melati mentioned that she had compared the system used in this program with other automated writing

evaluation program's systems and she admitted that "this program has the easiest access". Then, Bintang also felt "the use of the program was easy, especially for the downloading process and the utilization", since he did not face any barriers in downloading the result of the evaluation. With the ease of accessed offered by *Grammarly*, these students can feel the second beneficiary impact given in utilizing this program. They can easily monitor the feedback and correction given, and also faced no barriers in downloading the result of the evaluation. Though, in the interview, Mawar did not talk about the ease of access she felt during her experience on using this program. She did not mention any of her experience in having the ease of access nor the barriers in accessing her *Grammarly* account.

c. *High Rate of Evaluation Speed*

Another benefit found during experiencing *Grammarly* utilization is the high rate of evaluation speed. This benefit was felt by Mawar and Melati in their experiences. In the interview, Mawar mentioned that the process of checking is "quite quick" for her. Though, she also admitted that she needed "a longer time to read the result of the evaluation again". Similar answer to its positive impact on high speed of evaluation was also stated by Melati. She admitted that this program "could evaluate the writing quickly", but she was quite doubt on it and added that "the speed of evaluating process is depended on the availability of the connectivity to the internet". The satisfaction of having a high rate of evaluation speed can also become their consideration in having Grammarly as a tool in evaluating their academic writing. The high rate of evaluation speed can save their time in evaluating their academic writing and it can help them in revising their academic writing in a shorter time. Meanwhile, Bintang did not share any of his experience on having a high rate of evaluation speed nor talked about the duration of this program in evaluating his academic writing.

d. *Free Service*

The free service offered also becomes one of *Grammarly*'s strengths of utilizing this program. Mawar mentioned that she has been "satisfied with the evaluation given" by the Free-*Grammarly* since she viewed "it has been very helpful" for her. Moreover, the free service in checking writing offered by this program also made Melati "prefer to choose *Grammarly* than other AWE programs". Melati mentioned that she has felt satisfied with the free service given and even, she admitted that "the free-service given by *Grammarly* is better than other automated writing evaluation programs' services". Meanwhile, Bintang only mentioned that "the free service is the reason for choosing *Grammarly*" in evaluating his academic writing. Even though this program has a premium service, but these students seemed to prefer to choose the free-service one and has been satisfied with the result.

2. Weaknesses

These students also admitted that they faced several weaknesses in utilizing *Grammarly* in evaluating their academic writing, including misleading feedback, over-checking on reference list, and inability in checking context and content of writing.

a. *Misleading Feedback*

Even though *Grammarly* feedback is viewed as a useful correction for writing improvement, but some misleading feedbacks were also experienced by these students. In the interview, Mawar told me that the misleading feedback she experienced "changed her intentional meaning in the sentence" and it made her aware of the weaknesses of this program. This inaccuracy of evaluation result mentioned by her revealed a little of her dissatisfaction with the evaluation process and her less trust toward the evaluation given. Though, she still admitted that the error found might be caused by "her own fault".

A similar case was also experienced by Melati. She also shared some of her unexpected misleading feedbacks given by *Grammarly* which she considered it as the weaknesses she felt from utilizing this program. On her experience, the misleading feedbacks given by this program "was not in line with her intention" and she seemed to blame the system used by the program which gave her an inaccurate evaluation. She admitted that this problem "could mislead the context" of what she intended to state in her writing.

the teachers. An issue arises that the teachers are unfamiliar with the application of peer-assessment (Fazel, 2015) since they are having a lack of ability in implementing the peer-assessment in the classroom and having less sufficient training on affective aspect assessment. Furthermore, another issue appears that the teachers are having no trust (Tišma, 2016) in letting students doing the assessment caused of the incapability of students in rating their peers objectively and adequately (Saito & Fujita, 2004). Besides, the validity and reliability of peer-assessment are still questionable regarding its implementation in the classroom which leads to misunderstanding between teacher and students (Lan, Sung, & Chang, 2009).

Otherwise, different studies have investigated the implementation of peer-assessment and revealed different perspectives on its utilization. An investigation of peer-assessment for motivational aspect of 42 English department university students investigated by Fazel (2015) revealed that the students' perceptions about peer assessment are both

---

Possibly confused word

affective → effective

The word *affective* doesn't seem to fit this context. Consider replacing it with a different one.

Incorrect: *Jordan doesn't remember were he parked his car.*

Correct: *Jordan doesn't remember where he parked his car.*

Incorrect: *Which sign is usually paced on street corners to warn you of cross traffic?*

Correct: *Which sign is usually placed on street corners to warn you of cross traffic?*

LESS   IGNORE

Figure 2. Misleading feedback found on Melati's writing

Bintang also experienced the same case on misleading feedback. In his case, Bintang shared that the correction given led him to his confusion in his writing. He told me that "the feedback given was not the same as his idea" and it led him to his confusion. Even himself as the writer could not catch the purpose of the feedback since "it was out of his intention". As the result, he let his writing as it was written before and did not change it.
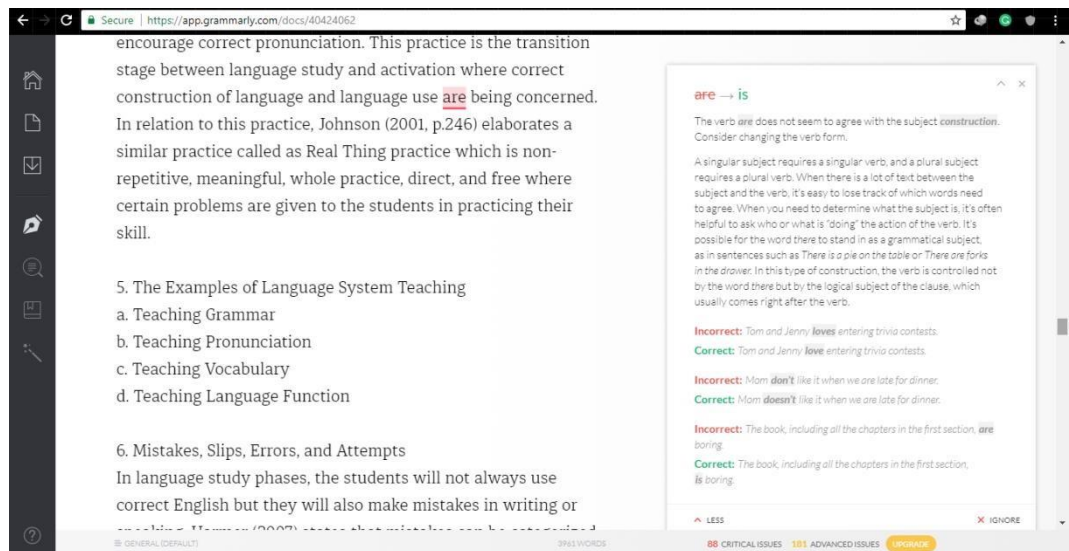

Figure 3. Misleading feedback found on Bintang's writing

Some of the misleading feedbacks experienced by these students reflected how this program can give an inaccuracy result on its evaluation. The feedback was not in line with the user's intention and even changed the intended meaning and emphasis of the sentence they made.

c. *Inability in Checking Context and Content of Writing*

Another weakness experienced on this automated writing evaluation is the inability in checking context and content of the writing. Melati told me that her intention to send a piece of writing to be proofread was to check not only on grammatical and mechanical aspect of writing, but also on coherency and cohesiveness of the writing, "whether the idea has been connected well or it has no relation to each other". However, *Grammarly* seemed has not fulfilled her satisfaction. Then, Bintang also shared his expectation that "context detection is needed as one feature of an automated writing evaluation program" and then, he realized that this program has not had the feature in detecting the relation

between one sentences to another. Thus, he admitted that this program has its weakness on detecting the context of writing. Meanwhile, Mawar did not share any of her experience regarding this issue. She confessed that she utilized *Grammarly* to check her grammar only, not to check the content of her writing.

Discussion

This main goal of this study was to reveal the strengths and weaknesses of using *Grammarly* in evaluating academic writing by exploring Indonesian EFL postgraduate students' experiences. In line with the goal, the strengths and weaknesses of this program has been revealed through the interview.

Some aspects as the strengths of this program are viewed to bring positive contributions for these students. By having feedback completed with a brief explanation and example, they can learn independently and do their self-editing and self-revising on their writing (Chen & Cheng, 2008). Then, with the ease of access offered by *Grammarly*, they can easily monitor the feedback and the correction given by this program, since they can have their full-access of their account (Chen & Cheng, 2008). Moreover, the satisfaction of having a high rate of evaluation speed expressed by these students also showed the feature of this program in saving time in evaluating their academic writing (Chou, et al., 2016; Cotos, 2011; Roscoe, et al., 2017) and it can help

them in revising their academic writing in a shorter time. Indirectly, these students also revealed that the free-service given by *Grammarly* can save their cost in proofreading their academic writing (Chou, et al., 2016). The strengths revealed in *Grammarly* can be classified as essential features had by a AWE program, since other AWE programs are also having similar features which support their user to have a better quality in writing.

Meanwhile, as coherency and cohesiveness of the text is a major concern in writing (Chen & Cheng, 2008; Wang & Wang, 2012), having

connected context and content become a matter of having a good quality of writing. However, the service given in free service of *Grammarly* has not fulfilled these matters and as the result, it

was not in line with these students' expectation on using this program as tools in proofreading. Even though, an additional optional service of proofreading offered in the Premium subscription, the service does not include any process of reorganizing, restructuring, or rewriting the content of writing since it is claimed that "issues regarding content, style, or voice are at the discretion of the author" (Grammarly, 2017).

In addition, the differences of strengths and weaknesses experienced by Mawar, Melati, and Bintang can be influenced by their length of time in utilizing this program, English type used in writing, and length of writing submitted. First, as Mawar has the shortest experience length of this program utilization, only seven months, it indicated that she has not experienced any specific problems in utilizing this program yet. Indifferent from Melati and Bintang who have longer experiences, Melati for one year and a half and Bintang for two years and a half, they have applied this program more frequent and felt more benefits and drawbacks given by the program than Mawar. Second, Mawar used American English on her writing which is different from Melati and Bintang who used British English. Thus, Mawar did not face any conflict with the *Grammarly* system which labeled any other type of English as non-American variant. Third, Mawar only submitted one to two paragraphs of her writing to Grammarly while Melati and Bintang submitted a whole full paper, including the reference list. The less writing submitted by Mawar cannot reveal further strengths and weaknesses of the program like Melati and Bintang's writing which can reveal the weaknesses on over- checking the reference list and checking the context and content of writing. Thus, the length of experience, type of English used, and length of writing submitted can influence the result of this program evaluation and also the perspective of the user in utilizing this program.

CONCLUSION

The current study helps to reveal the strengths and weaknesses of *Grammarly* as an automated writing evaluation program. The experiences shared by these students can be valuable considerations in deciding the use of *Grammarly* in evaluating academic writing. With the ease of access, high rate of evaluation speed, and free access given by *Grammarly*, it can help the users to proofread their academic writing without

spending time and more cost budget. The useful color-coded feedback completed with a brief explanation and simple example can help them in understanding the errors and make them learn better. Though, the caveats found, such as several misleading feedbacks, weaknesses on detecting the type of English and reference list, and lack of context and content evaluation experienced, can also be a challenge for them in utilizing this program. Thus, these strengths and weaknesses revealed can be a consideration for any independent author or writer in deciding to utilize this program in evaluating their academic writing.

From the result, it can be seen that *Grammarly* still has some limitations and weaknesses in giving the service. However, with a combination of *Grammarly*'s strengths and students' background knowledge on writing, a better quality of writing might be achieved. Thus, it is recommended for other researchers to investigate the efficiency of the feedback given by *Grammarly* in improving students' writing quality. Moreover, as this study focused on using a Free-version of *Grammarly*, a limited service may be experienced by these students, and thus, further study in investigating the Premium-access of *Grammarly* is needed to seek a whole full service given by this program to achieve better understanding of this program's strengths and weaknesses.

REFERENCES

Chen, C. F. E., & Cheng, W. Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology, 12*(2), 94-112. Retrieved from https://www.researchgate.net/publication/45681611

Chou, H. C., Moslehpour, M., & Yang, C. Y. (2016). My access and writing error corrections of EFL college pre-intermediate students. *International Journal of Education, 8*(1), 144-161. Retrieved from https://www.researchgate.net/publication/299535444

Cotos, E. (2011). Potential of automated writing evaluation feedback. *CALICO Journal, 28*(2), 420-459. Retrieved from https://www.researchgate.net/publication/274630246

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson Education.

Ebyary, K. E., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies, 10*(2), 121-142. Retrieved from http://files.eric.ed.gov/fulltext/EJ936915.pdf

Ferster, B, Hammond, T. C., Alexander, C., & Lyman, H. (2012). Automated formative assessment as a tool to scaffold student documentary writing. *Journal of Interactive Learning Research, 23*(1), 21-39. Retrieved from https://eric.ed.gov/?id=EJ979112

Grammarly. (2017). *About Grammarly.* Retrieved from https://support.grammarly.com/hc/en-us/categories/115000018611-About- Grammarly

Roscoe, R. D., Wilson, J., Johnson, A. C., & Mayra, C. R. (2017). Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation. *Computers in Human Behavior, 70*, 207-221. doi: 10.1016/j.chb.2016.12.076.

Scharber, C., Dexter, S., Riedel, E. (2008). Students' experiences with an automated essay scorer. The Journal of Technology, *Learning, and Assessment, 7*(1), 1-45. Retrieved from http://www.jtla.org

Wang, F. & Wang, S. (2012). A comparative study on the influence of automated evaluation system and teacher grading on students' English writing. *Procedia Engineering, 29*(2012), 993-997. doi:10.1016/j.proeng.2012.01.077

Wang, P. L. (2013). Can automated writing evaluation programs help students improve their English writing? *International Journal of Applied Linguistics & English Literature, 2*(1), 6-12. doi:10.7575/ijalel.v.2n.1p.6

Wilson, J., & Andrada, G. N. (2016). Using automated feedback to improve writing quality: opportunities and challenges. In Rosen, Y., Ferrara, S., Mosharraf, M. (eds). (2016). *Handbook of Research on Technology Tools for Real-World Skill Development* (pp. 678-703). IGI Global: US. doi: 10.4018/978-1-4666-9441-5.ch026

Wilson, J. (2016). Associated effects of automated essay evaluation software on growth in writing quality for students with and without disabilities. *Reading and Writing, 30*(4), 691-718. doi: 10.1007/s11145-016-9695-z